Arabic Sentiment Analysis using Deep Learning for COVID-19 Twitter Data

Sarah Alhumoud

Computer Science Department, Al Imam Mohammad Ibn Saud Islamic University, (IMSIU), Saudi Arabia

Abstract

Novel coronavirus, (COVID-19) first noticed in December 2019, and became a world pandemic affecting not only the health sector, but economic, social and psychological wellbeing as well. Individuals are using social media platforms to communicate feelings and sentiments on this pandemic. This article aims at analyzing and visualizing the influence of coronavirus (COVID-19) using machine learning and deep learning methods to quantify the sentiment shared publicly corelated with the actual number of cases reported over time. On the analysis of 10 Million Arabic tweets, results show that deep learning techniques using an ensemble model outperformed machine learning using SVM with an accuracy of 90% and 77% respectively. It also outperformed the individual deep learning models.

Key words:
COVID-19, machine learning, sentiment analysis, social computing.

1. Introduction

Coronaviruses emergence as pathogenic was in 2003 and 2012 with the appearance of the severe acute respiratory syndrome in China followed by the Middle East respiratory syndrome in Saudi Arabia [1],[2]. According to the world health organization (WHO) situation report, In December 2019, a series of patients with symptoms of unknown cause emerged in Wuhan, China leading to the world-wide pandemic of COVID-19 [3].

Now and according to WHO's report on the 9th of Sept 2020 that there are more than 27M confirmed cases with around 1M deaths caused by COVID-19 around the world. In Saudi Arabia the number of confirmed cases on the same date mentioned exceeded 300K with more than 4k deaths. This pandemic not only affected health, but also socioeconomics on the large and small scale, globally and personally. This is reflected greatly with the conversation exchange in Twitter. Twitter is used heavily in Saudi Arabia for communicating news by authorities and government parties. Also, by individuals to share opinions, that may lead to social change. In 2019 more than half of the Saudi population used Twitter [4], and several occasions proved that the public buzz and talk in Twitter is indeed leading to change and improvements in services in the Saudi context. The contribution of this study is as follows: first, collect a corpus of 10M tweets related to the COVID-19 topic in Arabic. Second, the preprocessing of the collected data and having as a result 416,292 tweets. Third, is implementing the machine learning and deep learning models to analyze data over time. Forth, is the visualization of the classification results as sentiments. Fifth. is the visualization of the health reports pulled from the Center of System Science and Engineering (CSSE) at Johns Hopkins University [49] Github repository.

2. Related Work

The related work is divided into two sections. The COVID-19 related literature, and the sentiment analysis using deep learning section. We start with former here.

2.1 COVID-19 Data Analysis Literature

The literature concerned with the analysis of COVID-19 data was considered from January 2020 until July 2020. Studies are mainly aiming to do four of the following tasks, first, social media analysis including network graph analysis, and social network text analysis. Second, chest x-ray image processing using machine learning and deep learning to predict cases. Third, cellular data analysis that corelates the handover volume along with moving patterns with the forecasted risks of infection. Finally, general cases predictions and forecast for potential numbers of infected cases from the confirmed and death cases reports. Those types of studies are summarized in Table 1. In this study the focus is on the first type of analysis, the social media analysis and sentiment analysis in specific. This is presented in the next subsection.

Manuscript received September 5, 2020
Manuscript revised September 20, 2020

DOI: 10.22937/IJCSNS.2020.20.09.16

Table 1: literature on COVID-19 analysis

Type of Analysis	*References*
Social media analysis	[3] , [5], [6], [7], [8], [9], [10], [11], [12], [13, p. 19], [14]
Chest x-ray image processing	[15], [16],[17],[18], [19],[20],[21]
Cellular data analysis	[22],[23, p. 19]
Cases predictions	[24],[25],[26],[27],[28, p. 19], [29],

2.1.1 Social media analysis

Mining data for meaning and trends from social media is a low-cost process that yields insight. COVID-19 related data analysis in social media is concerned with three main platforms, Twitter, Chinese's Weibo, and Facebook. Twitter analysis was either topic modeling [6], [7], [3], [12], [14], or sentiment analysis [9], [13, p. 19], or location mapping [10], [12]. The following is a brief review to the studies for each social platform.

Sear et. al. [5] proposed a machine learning model to analyze and understand misinformation around COVID-19 by analyzing posts and replies in Facebook for the first two months of 2020. The 8,277 documents were collected and annotated manually, then analyzed according to the coherence scoring approach into topics using Latent Dirichlet Allocation (LDA) [30]. The study concluded that the coherence score C_v for pro-vaxin communities is larger than that for anti-vaxin, indicating a more focused discussion around COVID-19 by the former community. However, it is noticed that the C_v scores were irregular showing an oscillating behavior with no clear trend or correlation to the number of topics. This needs an extra investigation in terms of dataset size and selected algorithm parameters.

Zhao et al. [6] analyzed the sentiment of search topics around COVID-19 in the Chinese Sina Microblog platform, Weibo for the first two months of 2020. They used the ROST CM6.0 sentiment analysis tool and obtained 7 sentiments, 3 levels of both positive and negative; and neutral, no accuracy reported.

Muthusami et al. [7] collected Tweets on the last 2 weeks of March 2020. Tweets were classified using an ensembled classifier model, LogitBoost, that gave comparable results to the Naive Bayes classifier scoring 74% and 72% in accuracy respectively.

Abd-Alrazaq et al. [3] collected tweets on February and first 2 weeks of March of 2020. Initially collected 2,787,247 tweets, and on removing non-English and duplicates, the remaining was 6% of the tweets, 167,073 tweets. The analysis involved the identification of 12 topics and 4 themes in an effort to identify the effect of fake news.

Li et al. [8] 31 analyzed Weibo posts one week before the COVID-19 declaration and one week after in the interval 13–26 January, 2020. They Used Online Ecological Recognition (OER) [31] to acquire the psychological states. The study calculated word frequency, scores of emotional indicators (e.g., anxiety, depression, indignation, and Oxford happiness) and cognitive indicators (e.g., social risk judgment and life satisfaction) from the collected data. The sentiment analysis and the paired sample t-test were performed to examine the differences in the same group before and after the declaration of COVID-19 on 20 January.

Manguri et al. [9] collected tweets for 7 consecutive days from April 9th 2020. The aim of the study is to test tweet sentiment and the subjectivity in terms of testing whether a tweet expresses a fact or a feeling. Feelings are tested for 10 different emotions, and those are: happy, confident, optimistic, hopeful, calm, neutral, relieved, pessimistic, worried, discouraged, depressed. Although the authors used Python for data collection and analysis However, the results are reported while not identifying accuracy indicators or classification algorithm used.

Jahanbin et al. [10] collected tweets from twitter and analyzed them using fuzzy rule-based evolutionary algorithm called Eclass1-MIMO. The results were mapping tweets to the word map.

J et al. [11] collected Chinese messages from Weibo for 39 days starting from 23 Dec 2019. The study did 2 analysis quantitive and qualitative. For the former, the aim was to determine if Weibo messages were an indicator to the number of cases reported. Using a linear regression model the study displays that there is a positive correlation between the number of posts and the number of reported cases. For the latter, 3 themes were manually detected in the messages: knowledge, beliefs, and health behaviors.

Singh et al. [12] collected tweets for 2 months starting from 15 Jan 2020. They find out a relation between tweet volume and the number reported cases. This was done through analyzing conversations mentioning country names to infer the location. The study identified a positive correlation between them, and that conversations are corelated with reported cases with a lead rather than a lag. also, this study carried out a topic modeling analysis and grouped tweets into 8 different topics. Those are: economy, emotion, healthcare, global Nature, information providers, social, government, individual concerns. Moreover, the study analyzed the spread of 5 different rumors in terms of volume. The study did not mention the algorithms used for modeling or classification.

Mathur et al. [13, p. 19] collected 30,000 tweets using TweetBinder [32] for 83 days starting from 22 Jan 2020. They analyzed data using Word-Emotion Association Lexicon (EmoLex) [33]. It is a list of English words and their association to 8 different emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and 2 sentiments (positive, negative). The analysis was carried out using machine learning with no specific algorithm mentioned with 80% accuracy.

Santis et al. [14] proposed a framework for topic detection and tracking exploring terms through time using graph analysis. The dataset is collected tweets containing 35 Italian keywords for 77 days starting from 9 March 2020. The authors claim that the proposed system could be generalized to be used for any language to detect and track topics emerging from a specific event.

Table 2: Literature on social network analysis for COVID-19

Reference	*Platform*	*Type of Analysis*	*Days Collected*	*Dataset size (P= post, T=tweet, U= user)*
[5]	Facebook	LDA Topic modeling	60	8,277P
[6]	Weibo	SA for Search topics	60	$5x10^8$P
[7]	Twitter	SA for tweets	14	18,216T
[3]	Twitter	LDA Topic modeling	14	167,073T
[8]	Weibo	SA for user posts	14	17,865U
[9]	Twitter	SA for tweets	9	530,232T
[10]	Twitter	Location mapping	-	364,080T
[11]	Weibo	Analyze messages	39	115,299P
[12]	Twitter	Location mapping, topic modeling	60	2,792,513T
[13, p. 19]	Twitter	Emotion detection	83	30,000T
[14]	Twitter	Topic modeling	77	1,044,645T

2.1.2 Deep learning methods for sentiment analysis

Deep learning is a branch of machine learning based on artificial neural networks. it has proven its success over classical machine learning techniques in multiple domains including computer vision, speech recognition, natural language processing (NLP) and sentiment analysis [34] [35]. Sentiment analysis is an NLP task that aims at classifying words, paragraphs or documents according to their polarity into positive or negative [36]. Among different machine learning techniques SVM proved to be superior for the task of Arabic sentiment analysis [37], [38]. With recent advent in deep learning, different architectures are used for sentiment analysis tasks such as, convolutional neural networks (CNNs) [39], and recurrent neural networks (RNNs) [40]. RNNs are among the oldest neural networks' techniques but they suffer from the vanishing and exploding gradient problem [41] where the gradient values either grow only larger to the point of explosion or shrink smaller to the point of vanishing. A resolution to this issue, is adding a gaiting mechanism with what is called gated RNN or gated recurrent units (GRUs) [40],[42] and long short-term memory (LSTM) [43],[44]. In this study we will be implementing a stacked GRU (SGRU) [45] and a stacked bi-GRU (SBGRU) [46] along with an ensemble model based on those models along with AraBERT [47]. AraBERT is an Arabic pretrained language model based on Google's BERT architecture that was used to classify data.

3. Methods

Sentiment analysis passes through multiple steps starting from data collection, followed by data preprocessing, then classification. Those steps are described in the following and depicted in figure 1.

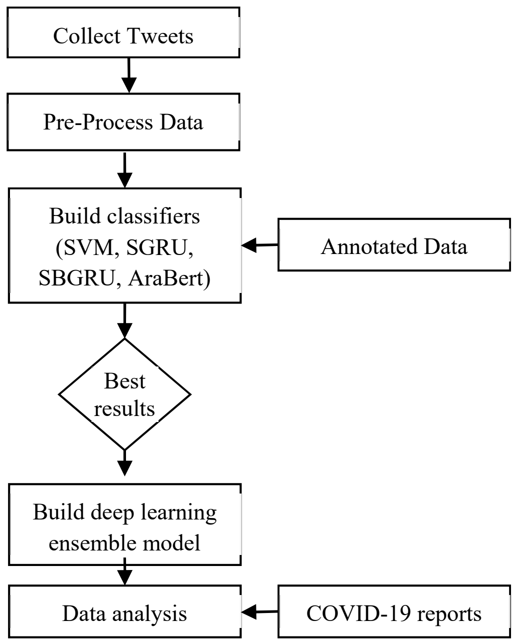

Figure 1: the study architecture with ensemble deep learning model and data analysis

3.1 Dataset Description

Using Twitter streaming API and "twitterR" package for R, a total of 10M tweets were collected in April 2020. Tweets are collected using 7 keywords: ("كورونا" ,"كرونا", "covid", "فايرس" ,"كارونا" , "فيروس", "corona"). The keywords are in Arabic and English, but the dataset is filtered to keep Arabic tweets only.

3.2 Data Preprocessing

Data preprocessing is necessary to minimize the noise in data and maximize the classification accuracy and success. The first step in data preprocessing is the spam filtration. Using the algorithm provided by [48], data was deducted by around 95% left with 416,292 tweets only. Spam filtering include deleting duplicate tweets or tweets that satisfies one of the following conditions. Either having one of the predefined spam keywords or having more than 4 hashtags. As advertisement tweets tent to have those two features. The second step in data preprocessing is data cleaning. This is done by deleting stop words, deleting all non-Arabic letters from each tweet, and more than 2 duplicates of each letter in a word. The third and last step is normalization, this is done by unifying multiple-form letters into one of its forms [38] an example to that is unifying (أ،إ،آ،ا) into (ا).

3.3 Classification

Data classification was implemented in several steps: first is to build and run SVM classifier to compare the results with the deep learning results. Then implementing the deep learning models, SGRU and SBGRU along with the newly available AraBert classifier. Those three models are then used to build the ensemble model by voting the best accuracy among them, the results in the following section show that the accuracy of the ensemble model outperform every other individual deep learning model under consideration.

4. Results and Discussion

The analysis was carried out on the daily reports of confirmed cases published by the (CSSE) at Johns Hopkins University [49]. This data is coupled with the analysis results of the deep learning model depicted in Figure 1. On analyzing the most frequent words that were used in the data collected, "كورونا" which means "corona" scored the highest with a frequency of 221,455 words. Although different keywords are used to express the virus, this was the most dominant. immediately after that is the words "فايرس" and "الله" meaning "virus" and "God" respectively. This reflects the spiritual influence in people's communication and expression. The rest of the top words are plotted in Figure 2. Figure 3 shows confirmed and death cases starting from February until the date of writing this paper in September, pulled from [49] GitHub. This figure shows that the trend of confirmed cases is starting to flatten by September. Figure 4 shows the amount of different sentiments on applying the deep learning ensemble model to the original dataset. The results show that the negative sentiment is dominant as opposed to the positive and neutral. This is expected as users are more inclined to express negative feelings more than positive ones [37], [50]. Also, the number of tweets was corelated to major country rules and legislations. On the 6th of April the 24h quarantine rule started in most of the cities, leading to major explosion in the communication exchange on the topic of COVID-19 in Twitter as shown by point (A) in Figure 4.

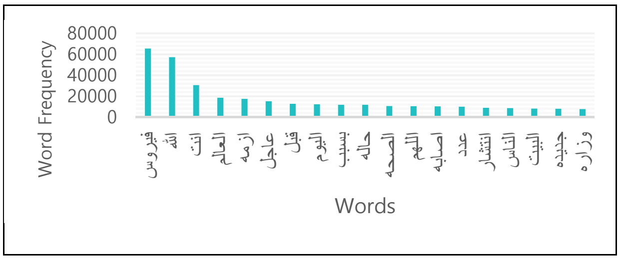

Figure 1: Word frequency for the top 19 words

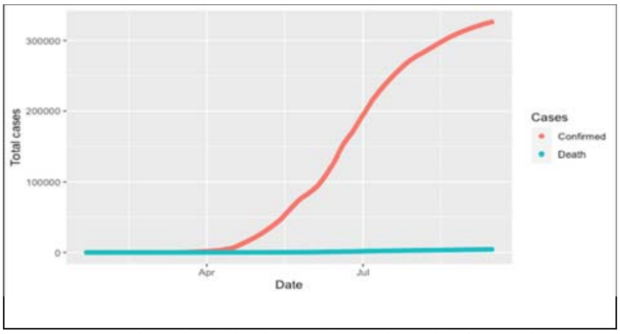

Figure 2: Confirmed and death cases in Saudi Arabia form Feb to Sept 2020

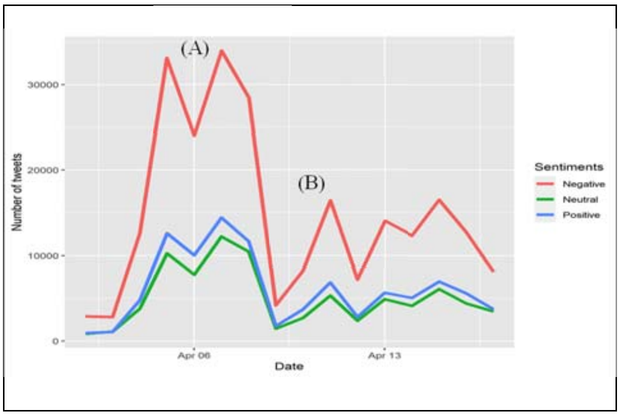

Figure 4: Tweets' sentiment classification

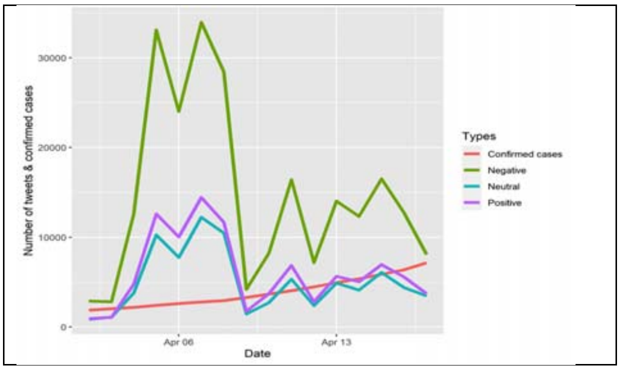

Figure 5: Tweets' sentiment classification & confirmed cases in April 2020

Table 3: Results of individual and ensemble model

Model	*Accuracy*
SVM	77.92%
SGRU-6	82.08%
SBIGRU-5 l	81.59%
AraBERT	85.41%
Ensemble	90.21%

On the 11th of April the 24h quarantine rule was confirmed to be extended and this is related to the rising trend of tweets reflected in point (B), Figure 4. Figure 5 presents a joint plot of confirmed cases and the sentiment in the month of April showing the rise from only several thousand cases to reach around 300k cases in September 2020.
To evaluate the performance of the different models we use accuracy which is the percentage of correctly classified tweets to the total number of tweets as in Equation 1.

$$Accuracy = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}) \quad (1)$$

The performance of the SGRU and SBGRU was superior in levels 6 and 5 respectively scoring 82.08% and 81.59% respectively, Table 3. Therefore, the performance of those levels along with AraBERT is used to build the ensemble model by voting the best accuracy among all three algorithms. The ensemble model accuracy was the highest with an accuracy of 90.21%. The results of all individual models along with the ensemble model is depicted in Table 3. As shown in the table the machine learning model SVM scored the lowest accuracy among all considered models with 77.92%.

5. Conclusions

This study aimed at studying the publicly available Twitter tweets in relation to the COVID-19 data. 10M tweets were collected in April 2020. Those tweets where analyzed using both machine learning and deep learning models to classify the data's sentiment. Results show that the best performance was achieved by the ensemble model of 3 different deep learning models, SGRU, SBGRU, and AraBERT. As a continuation to this work we plan to analyze more data by adding more processing resources.

References

[1] A. Zumla, J. F. W. Chan, E. I. Azhar, D. S. C. Hui, and K.-Y. Yuen, "Coronaviruses - drug discovery and therapeutic options," *Nat. Rev. Drug Discov.*, vol. 15, no. 5, p. 327—347, May 2016, doi: 10.1038/nrd.2015.37.

[2] J. Cui, F. Li, and Z.-L. Shi, "Origin and evolution of pathogenic coronaviruses," *Nat. Rev. Microbiol.*, vol. 17, no. 3, p. 181—192, Mar. 2019, doi: 10.1038/s41579-018-0118-9.

[3] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, and Z. Shah, "Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study," *J Med Internet Res*, vol. 22, no. 4, p. e19016, Apr. 2020, doi: 10.2196/19016.

[4] A. AlFarasani, T. AlHarthi, and S. AlHumoud, "ATAM: Arabic Traffic Analysis Model for Twitter," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 3, 2019, doi: 10.14569/IJACSA.2019.0100343.

[5] R. F. Sear *et al.*, "Quantifying COVID-19 Content in the Online Health Opinion War Using Machine Learning," *IEEE Access*, vol. 8, pp. 91886–91893, 2020.

[6] Y. Zhao, S. Cheng, X. Yu, and H. Xu, "Chinese public's attention to the COVID-19 epidemic on social media: Observational descriptive study," *J. Med. Internet Res.*, vol. 22, no. 5, 2020, doi: 10.2196/18825.

[7] Muthusami R., Bharathi A, and Saritha K, "Covid-19 outbreak: Tweet based analysis and visualization towards the influence of coronavirus in the world," *GEDRAG Organ. Rev.*, vol. 33, no. 2, pp. 534–549, 2020.

[8] S. Li, Y. Wang, J. Xue, N. Zhao, and T. Zhu, "The Impact of COVID-19 Epidemic Declaration on Psychological Consequences: A Study on Active Weibo Users," *Int. J. Environ. Res. Public. Health*, vol. 17, no. 6, p. 2032, Mar. 2020, doi: 10.3390/ijerph17062032.

[9] Kamaran H. Manguri, Rebaz N. Ramadhan, and Pshko R. Mohammed Amin, "Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks," *Kurd. J. Appl. Res.*, vol. 5, no. 3, pp. 54–65, May 2020, doi: http://doi.org/10.24017/kjar.

[10] Kia Jahanbin and Vahid Rahmanian, "Using twitter and web news mining to predict COVID-19 outbreak," *Asian Pac. J. Trop. Med.*, vol. Manuscript submitted for publication, 2020, doi: 10.4103/1995-7645.279651.

[11] J. Li, Q. Xu, R. Cuomo, V. Purushothaman, and T. Mackey, "Data Mining and Content Analysis of the Chinese Social Media Platform Weibo During the Early COVID-19 Outbreak: Retrospective Observational Infoveillance Study," *JMIR Public Health Surveill.*, vol. 6, no. 2, p. e18700, Apr. 2020, doi: 10.2196/18700.

[12] L. Singh *et al.*, *A first look at COVID-19 information and misinformation sharing on Twitter*. 2020.

[13] A. Mathur, P. Kubde, and S. Vaidya, "Emotional Analysis using Twitter Data during Pandemic Situation: COVID-19," in *5th International Conference on Communication and Electronics Systems (ICCES)*, 2020, pp. 845–848.
[14] E. D. Santis, A. Martino, and A. Rizzi, "An Infoveillance System for Detecting and Tracking Relevant Topics From Italian Tweets During the COVID-19 Event," *IEEE Access*, vol. 8, pp. 132527–132538, 2020.
[15] M. Abdel-Basset, R. Mohamed, M. Elhoseny, R. K. Chakrabortty, and M. Ryan, "A hybrid COVID-19 detection model using an improved marine predators algorithm and a ranking-based diversity reduction strategy," *IEEE Access*, pp. 1–1, 2020.
[16] M. A. Mohammed *et al.*, "Benchmarking Methodology for Selection of Optimal COVID-19 Diagnostic Model Based on Entropy and TOPSIS Methods," *IEEE Access*, pp. 1–1, 2020.
[17] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020.
[18] Z. U. Abideen *et al.*, "Uncertainty Assisted Robust Tuberculosis Identification With Bayesian Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 22812–22825, 2020.
[19] X. Ouyang *et al.*, "Dual-Sampling Attention Network for Diagnosis of COVID-19 From Community Acquired Pneumonia," *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2595–2605, 2020.
[20] Y. Pathak, P. K. Shukla, and K. V. Arya, "Deep bidirectional classification model for COVID-19 disease infected patients," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, pp. 1–1, 2020.
[21] C. Butt, J. Gill, D. Chun, and B. A. Babu, "Deep learning system to screen coronavirus disease 2019 pneumonia," *Appl. Intell.*, Apr. 2020, doi: 10.1007/s10489-020-01714-3.
[22] A. A. R. Alsaeedy and E. K. P. Chong, "Detecting Regions At Risk for Spreading COVID-19 Using Existing Cellular Wireless Network Functionalities," *IEEE Open J. Eng. Med. Biol.*, vol. 1, pp. 187–189, 2020.
[23] E. Hernández-Orallo, P. Manzoni, C. T. Calafate, and J. Cano, "Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: the case of COVID-19," *IEEE Access*, pp. 1–1, 2020.
[24] R. S. Yadav, "Data analysis of COVID-2019 epidemic using machine learning methods: a case study of India," *Int. J. Inf. Technol.*, May 2020, doi: 10.1007/s41870-020-00484-y.
[25] L. J. Muhammad, Md. M. Islam, S. S. Usman, and S. I. Ayon, "Predictive Data Mining Models for Novel Coronavirus (COVID-19) Infected Patients' Recovery," *SN Comput. Sci.*, vol. 1, no. 4, p. 206, Jun. 2020, doi: 10.1007/s42979-020-00216-w.
[26] K. Decock, K. Debackere, A.-M. Vandamme, and B. Van Looy, "Scenario-driven forecasting: modeling peaks and paths. Insights from the COVID-19 pandemic in Belgium," *Scientometrics*, vol. 124, no. 3, pp. 2703–2715, Sep. 2020, doi: 10.1007/s11192-020-03591-6.
[27] F. Rustam *et al.*, "COVID-19 Future Forecasting Using Supervised Machine Learning Models," *IEEE Access*, pp. 101489–101499, 2020.
[28] R. B. Duffey and E. Zio, "Analysing Recovery From Pandemics by Learning Theory: The Case of CoVid-19," *IEEE Access*, vol. 8, pp. 110789–110795, 2020.
[29] P. Staszkiewicz, I. Chomiak-Orsa, and I. Staszkiewicz, "Dynamics of the COVID-19 Contagion and Mortality: Country Factors, Social Media, and Market Response Evidence From a Global Panel Analysis," *IEEE Access*, vol. 8, pp. 106009–106022, 2020.
[30] S. Syed and M. Spruit, "Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation," in *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2017, pp. 165–174.
[31] M. Liu, J. Xue, N. Zhao, X. Wang, D. Jiao, and T. Zhu, "Using Social Media to Explore the Consequences of Domestic Violence on Mental Health," *J. Interpers. Violence*, vol. 0, no. 0, p. 0886260518757756, doi: 10.1177/0886260518757756.
[32] "TweetBinder." https://www.tweetbinder.com/.
[33] S. M. Mohammad and P. D. Turney, "Crowdsourcing a Word-Emotion Association Lexicon," *Comput. Intell.*, vol. 29, no. 3, pp. 436–465, 2013.
[34] Z. H. Hakami, M. Alshathri, N. Alqhtani, L. Alharthi, and S. AlHumoud, "Arabic Sentiment Analysis Using Deep Learning: A Review," 2019.
[35] W. Algihab, N. Alawwad, A. Aldawish, and S. AlHumoud, "Arabic Speech Recognition with Deep Learning: A Review," in *Social Computing and Social Media. Design, Human Behavior and Analytics*, vol. 11578, G. Meiselwitz, Ed. Cham: Springer International Publishing, 2019, pp. 15–31.
[36] S. Alhumoud, M. Altuwaijri, T. Albuhairi, and W. Alohaideb, "Survey on Arabic Sentiment Analysis

in Twitter," *Int. J. Comput. Inf. Eng.*, vol. 9, no. 1, pp. 364–368, 2015.

[37] S. Alhumoud, "Twitter Analysis for Intelligent Transportation," *Comput. J.*, vol. 61, no. 12, p. bxy129, 2018, doi: 10.1093/comjnl/bxy129.

[38] S. Alhumoud, T. Albuhairi, and W. Alohaideb, "Hybrid Sentiment Analyser for Arabic Tweets using R," in *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015)*, Lisbon,Purtogal, 2015, vol. 1.

[39] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1746–1751, doi: 10.3115/v1/D14-1181.

[40] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder–Decoder Approaches," in *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, Doha, Qatar, Oct. 2014, pp. 103–111, doi: 10.3115/v1/W14-4012.

[41] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *International conference on machine learning*, 2013, pp. 1310–1318.

[42] D. Tang, B. Qin, and T. Liu, "Document Modeling with Gated Recurrent Neural Network for Sentiment Classification," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, 2015, pp. 1422–1432, doi: 10.18653/v1/D15-1167.

[43] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[44] K. S. Tai, R. Socher, and C. D. Manning, "Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Beijing, China, 2015, pp. 1556–1566, doi: 10.3115/v1/P15-1150.

[45] D. Anil, A. Vembar, S. Hiriyannaiah, S. Gm, and K. Srinivasa, "Performance Analysis of Deep Learning Architectures for Recommendation Systems," in *2018 IEEE 25th International Conference on High Performance Computing Workshops (HiPCW)*, Bengaluru, India, Dec. 2018, pp. 129–136, doi: 10.1109/HiPCW.2018.8634192.

[46] D. Ma, S. Li, and H. Wang, "Target Extraction via Feature-Enriched Neural Networks Model," in *Natural Language Processing and Chinese Computing*, vol. 11108, M. Zhang, V. Ng, D. Zhao, S. Li, and H. Zan, Eds. Cham: Springer International Publishing, 2018, pp. 353–364.

[47] W. Antoun, F. Baly, and H. Hajj, "AraBERT: Transformer-based Model for Arabic Language Understanding," *ArXiv*, vol. abs/2003.00104, 2020.

[48] N. Al Twairesh, M. Al Tuwaijri, A. Al Moammar, and S. Al Humoud, "Arabic Spam Detection in Twitter," May 2016, pp. 38–43.

[49] *CSSE at Johns Hopkins University COVID-19 data*. .

[50] C. Collins, S. Hasan, and S. Ukkusuri, "A Novel Transit Rider Satisfaction Metric: Rider Sentiments Measured from Online Social Media Data," *J. Public Transp.*, vol. 16, no. 2, Jun. 2013, doi: http://dx.doi.org/10.5038/2375-0901.16.2.2.